6-12-2014

Transcription factor cooperativity in early adipogenic hotspots and super-enhancers

Rasmus Siersbæk
Syddansk Universitet

Atefeh Rabiee
Syddansk Universitet, arabiee@pacific.edu

Ronni Nielsen
Syddansk Universitet

Simone Sidoli
Syddansk Universitet

Sofie Traynor
Syddansk Universitet

See next page for additional authors

Recommended Citation

Siersbæk, R., Rabiee, A., Nielsen, R., Sidoli, S., Traynor, S., Loft, A., Poulsen, L. L., Rogowska-Wrzesinska, A., Jensen, O. N., & Mandrup, S. (2014). Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Reports, 7*(5), 1443–1455. DOI: 10.1016/j.celrep.2014.04.042
https://scholarlycommons.pacific.edu/phs-facarticles/449

Cell Reports
Article

Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers

Rasmus Siersbæk,[1,2] Atefeh Rabiee,[1,2] Ronni Nielsen,[1] Simone Sidoli,[1] Sofie Traynor,[1] Anne Loft,[1] Lars La Cour Poulsen,[1] Adelina Rogowska-Wrzesinska,[1] Ole N. Jensen,[1] and Susanne Mandrup[1,*]
[1]Department of Biochemistry and Molecular Biology, University of Southern Denmark, 5230 Odense M, Denmark
[2]Co-first author
*Correspondence: s.mandrup@bmb.sdu.dk
http://dx.doi.org/10.1016/j.celrep.2014.04.042

SUMMARY

It is becoming increasingly clear that transcription factors operate in complex networks through thousands of genomic binding sites, many of which bind several transcription factors. However, the extent and mechanisms of crosstalk between transcription factors at these hotspots remain unclear. Using a combination of advanced proteomics and genomics approaches, we identify ~12,000 transcription factor hotspots (~400 bp) in the early phase of adipogenesis, and we find evidence of both simultaneous and sequential binding of transcription factors at these regions. We demonstrate that hotspots are highly enriched in large super-enhancer regions (several kilobases), which drive the early adipogenic reprogramming of gene expression. Our results indicate that cooperativity between transcription factors at the level of hotspots as well as super-enhancers is very important for enhancer activity and transcriptional reprogramming. Thus, hotspots and super-enhancers constitute important regulatory hubs that serve to integrate external stimuli on chromatin.

INTRODUCTION

A number of genome-wide studies on transcription factor binding in multiple different cell systems have shown that many transcription factors tend to colocalize with other factors on chromatin (Biddie et al., 2011; Grøntved et al., 2013; Heinz et al., 2010; Hurtado et al., 2011; Lefterova et al., 2008; Nielsen et al., 2008), and transcription factor hotspots occupied by multiple factors have even been described in some cell types (Boergesen et al., 2012; Chen et al., 2008; Gerstein et al., 2012; He et al., 2011; Moorman et al., 2006; Siersbæk et al., 2011). However, the functional significance of this colocalization is currently unclear. In addition, Whyte et al. (2013) and Lovén et al. (2013) recently demonstrated the existence of super-enhancers, which are large genomic regions (several kilobases) containing clusters of closely spaced transcription factor binding regions. These large super-enhancers are characterized by very high levels of Mediator subunit 1 (MED1) binding and seem to regulate cell identity. The relation between hotspots and super-enhancers is currently unclear.

Adipocyte differentiation is a well-studied differentiation process, and many of the transcription factors acting in a sequential manner to activate this differentiation process have been described (Farmer, 2006; Lefterova and Lazar, 2009; Rosen and MacDougald, 2006; Siersbæk et al., 2012b). 3T3-L1 preadipocytes differentiate into mature adipocytes in a rather synchronous and efficient manner upon exposure to a cocktail of adipogenic inducers, and we previously demonstrated that this is associated with extensive reprogramming of the chromatin landscape within the first 4 hr of differentiation, as evidenced by the dynamic change in DNase I hypersensitive (DHS) site profiles (Siersbæk et al., 2011). This cell line therefore represents an ideal model system for studying transcription factor cooperativity on chromatin during reprogramming of the genome.

Here, we combined advanced genomics and proteomics techniques to obtain molecular insight into the interplay among transcription factors that drive the early adipogenic reprogramming of 3T3-L1 cells. We demonstrate extensive colocalization of transcription factors in hotspots and super-enhancers, and show that hotspots are highly enriched in super-enhancer regions. Furthermore, our work reveals extensive cooperativity between transcription factors at the level of hotspots as well as super-enhancers, and indicates that this cooperativity is very important for transcriptional reprogramming during differentiation.

RESULTS

A Combined Genomics and Proteomics Approach Identifies Key Members of the Early Adipogenic Transcription Factor Network

Based on chromatin immunoprecipitation sequencing (ChIP-seq) profiling, we previously reported ~1,000 transcription factor hotspots occupied by five transcription factors during early 3T3-L1 adipogenesis (Siersbæk et al., 2011). Here, to further investigate the extent of hotspot formation and characterize their composition, we undertook a combined genomics and proteomics approach (Figure 1). First, we performed coimmunoprecipitation (coIP) of C/EBPβ-associated proteins 4 hr following induction of 3T3-L1 differentiation (Figure S1A), i.e., the time point at which we previously demonstrated dramatic chromatin

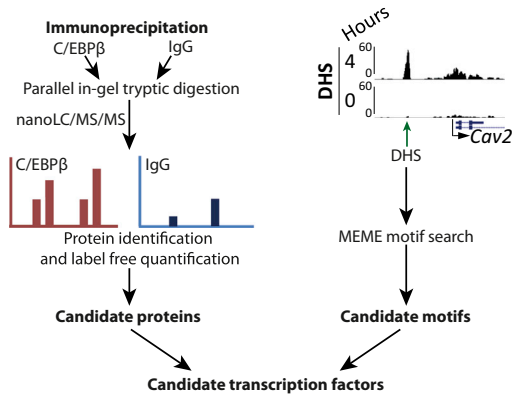


Figure 1. Schematic Overview of the Combined Genomics and Proteomics Approach Used to Identify Key Early Regulators of Adipocyte Differentiation

Motif analyses of DNA sequences at DHS sites 4 hr after induction of differentiation of 3T3-L1 cells obtained from previous analyses (Siersbæk et al., 2011) were combined with proteomics analyses of C/EBPβ-associated proteins to confidently identify candidate transcription factors involved in early adipogenic reprogramming.

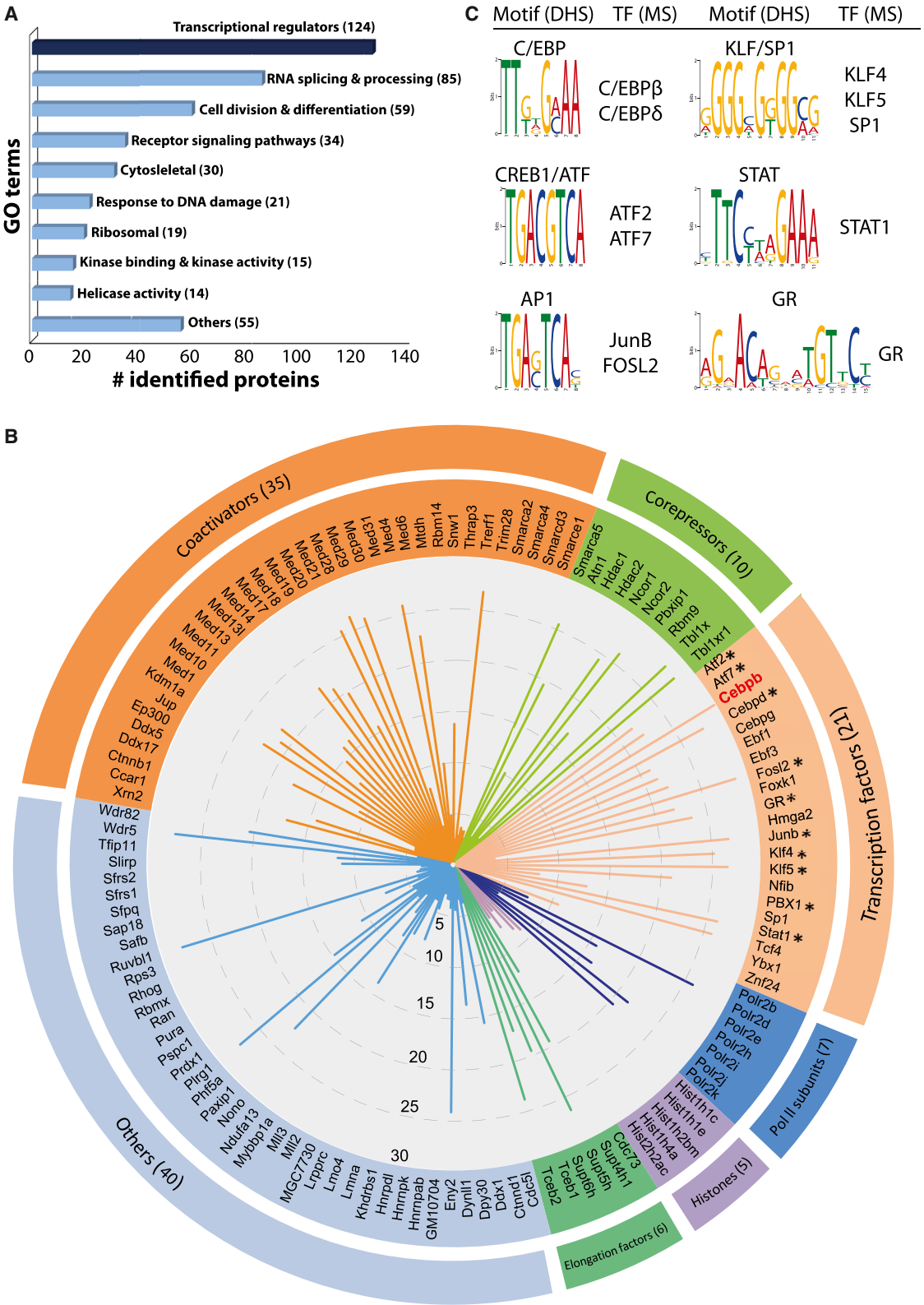
remodeling (Siersbæk et al., 2011). C/EBPβ was chosen as the bait for the proteomics analyses, because it has been shown to play an important role in regulating the early phase of adipocyte differentiation both in vitro and in vivo (Tanaka et al., 1997; Tang et al., 2003; Zhang et al., 2004), and it colocalizes extensively with the few factors we previously profiled by ChIP-seq (Siersbæk et al., 2011). The protein mixture was analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS; Figure 1, left).

We identified 292 proteins that coprecipitate with C/EBPβ in two independent biological replicates (Figure S1B; Tables S1 and S2). Gene Ontology (GO) annotation revealed that many of the identified proteins are transcriptional regulators, but we also identified proteins involved in RNA splicing and processing, as well as kinases, helicases, and ribosomal proteins (Figure 2A). The group of transcriptional regulators includes many coregulators and transcription factors (Figure 2B), some of which have previously been shown to associate with C/EBPβ, such as its heterodimerization partner C/EBPδ, as well as Krüppel-like factor 5 (KLF5) (Oishi et al., 2011), transcriptional intermediary factor 1β (TIF-1β) (Chang et al., 1998), and p300 (Mink et al., 1997). Most of these are highly enriched (>10-fold) in the C/EBPβ immunoprecipitation compared with the nonspecific immunoglobulin G (IgG) control, indicating a strong and specific association with C/EBPβ. Several of the transcription factors identified as C/EBPβ-interacting proteins by proteomics analysis have also been shown to regulate the early phase of adipocyte differentiation, e.g., KLF4 (Birsoy et al., 2008), KLF5 (Oishi et al., 2011), GR (Siersbæk et al., 2011; Steger et al., 2010), and PBX1 (Monteiro et al., 2011), clearly indicating that our approach is a powerful strategy for identifying biologically meaningful regulators of the differentiation process.

A comparison of the identified proteins with our previously published de novo motif analysis of DNA sequences at DHS regions identified at the 4 hr time point (Siersbæk et al., 2011; Figure S1A, right) revealed binding motifs for many of the transcription factors identified as C/EBPβ-associated proteins (Figure 2C). This indicates that these proteins bind directly to DNA at many open chromatin regions during early adipogenesis. Importantly, a major benefit of this combined approach is that it allows us to distinguish among different transcription factors that bind to the same motif, which is almost impossible based on sequence analyses alone. For example, we identify JunB and FOSL2 from the large AP1 family, and KLF4, KLF5, and SP1 from the large KLF/SP1 family as possible candidates for binding the AP1 and KLF/SP1 motif in DHS sites, respectively. Taken together, these results demonstrate the power of combining proteomics analyses of proteins associated with known key regulators with motif analyses of accessible chromatin regions in the genome to identify novel transcriptional regulators of biological processes.

Extensive Colocalization of Transcription Factors at Hotspot Regions

We chose to perform ChIP-seq profiling of eight factors from the combined proteomics and genomics screen described above (i.e., KLF4, KLF5, JunB, Fos-like antigen 2 [FOSL2], signal transducer and activator of transcription 1 (STAT1), activating transcription factor 2 [ATF2], ATF7, and PBX1 [indicated by asterisks in Figure 2B]) based on availability of high-quality antibodies. In addition, we chose to profile vitamin D receptor (VDR) (Blumberg et al., 2006; Cianferotti and Demay, 2007) and c-Jun (Mariani et al., 2007; Wang and Scott, 1994), which have previously been implicated in early stages of adipocyte differentiation. When combined with our previously published profiles of C/EBPβ, C/EBPδ, GR, and STAT5A (Siersbæk et al., 2011) as well as a new version of our previously published RXR profile, the results reveal a total of 54,724 transcription factor binding regions (~250–400 bp), most (58%) of which are occupied by more than one factor. Importantly, all of the investigated transcription factors colocalize with C/EBPβ on chromatin (see Figures 3B, 3C, and S2A), demonstrating a high degree of concordance between these genomics data and the proteomics analyses. Quantification of transcription factor colocalization shows that for all factors, most binding sites are occupied by additional factors, although the extent of colocalization with other factors seems to be factor dependent (Figure 3A). Importantly, the observed degree of transcription factor colocalization is much higher than that found for random sites (i.e., only 1.5% of randomized binding sites are occupied by more than one factor; Figure 3A, bottom). Based on this large number of investigated factors, we could identify ~12,000 hotspot regions that are occupied by at least five transcription factors, demonstrating that extensive colocalization of transcription factors is a common phenomenon. In fact, 40%–82% of the binding sites for a given factor are located in hotspots based on these data sets (Figure 3A), which is likely to be an underestimate, since we only analyzed a subset of the transcription factors that are active during this differentiation process. From the nine largest groups of hotspots, it is evident that many different types of hotspots are occupied by distinct subsets of factors (Figure 3B; the degree of co-occurrence of all transcription factor pairs at hotspots is illustrated in Figure S2A). Thus, it is unlikely that hotspots are the result of



(legend on next page)

unspecific associations between factors and accessible chromatin regions; instead, they are likely to be formed by the specific association of multiple factors with the same genomic regions. Interestingly, we identify 138 regions that are specifically targeted by all of the 15 investigated factors (Figure 3B). Given the importance of peroxisome proliferator-activated receptor γ (PPARγ) for adipocyte differentiation, it is highly interesting to note that two such major hotspots are located in close proximity to the *Pparg2* TSS (Figure 3C).

Given the high number of transcription factors that associate with hotspots, we performed re-ChIP experiments to assess whether transcription factors bind simultaneously or sequentially to hotspot regions (Figures 3D and S2B–S2E). For the sites investigated, we could demonstrate that five of the seven tested pairs of factors (i.e., JunB-C/EBPβ, KLF4-C/EBPβ, ATF2-JunB, ATF7-JunB, and KLF5-C/EBPβ) seem to occupy chromatin at these hotspots simultaneously, at least within the time resolution of the ChIP methodology. In contrast to these pairs, JunB and KLF4 show robust binding to the investigated regions in single ChIP experiments (Figure S2D), but they do not seem to occupy these regions at the same time, despite the fact that these factors are known to recognize completely different motifs (Figure 3D). The same is true for KLF4 and KLF5 (Figure S2C), which is expected because they bind to the same motif. Thus, in these data sets we find evidence for both simultaneous binding and dynamic sequential association of transcription factors with hotspots.

Hotspots Are Key Enhancer Regions

Analysis of the location of the identified transcription factor binding sites relative to genes revealed that even though all types of binding sites, in particular those occupied by few of the investigated factors, are enriched in gene promoters compared with a random control, most binding sites are found distal to transcription start sites, and this trend becomes more pronounced the more factors are bound to these regions (Figure 4A). Thus, hotspots are primarily found at gene distal regions. Interestingly, genome-wide profiling of the histone marks H3K4me1, H3K4me2, and H3K27ac, which were previously shown to characterize enhancers (Creyghton et al., 2010; Lupien et al., 2008; Rada-Iglesias et al., 2011), demonstrated that both distal non-hotspots (i.e., regions occupied by one to four factors) and hotspots (≥5 factors) are enriched for these three marks (Figure 4B). However, hotspots are associated with significantly higher levels of all marks than non-hotspot regions, demonstrating that many of the identified distal binding regions, and distal hotspots in particular, have the epigenomic profile of active enhancers.

Interestingly, genome-wide profiling of the coactivators Mediator subunit 1 (MED1), the histone acetyltransferase p300, and the SWI/SNF chromatin remodeling factor Brahma-related gene 1 (BRG1), reveals that the more factors that colocalize to a given region, the higher are the levels of coactivator recruitment (Figure 4C). Consistent with the increased BRG1 recruitment, the DHS-seq signal (Siersbæk et al., 2014, this issue of *Cell Reports*) also increases with the number of factors (Figure S3A). Importantly, the input control does not show a similar increase in signal compared with the coactivator ChIPs (Figure S3B). Taken together, these results indicate that transcription factors cooperate extensively at hotspots to remodel the chromatin, recruit coregulators associated with enhancer function, and establish an epigenomic enhancer profile.

To correlate the different types of transcription factor binding regions with transcriptional changes during the first 4 hr of differentiation, we employed 4-thiouridine (4sU)-RNA-seq, which is a robust and reproducible method that primarily maps newly synthesized RNA (Rabani et al., 2011; Figure S3C). Using this method, we identify 2,374 and 2,022 genes that are induced and repressed, respectively, during the first 4 hr of differentiation (Figure 4D). We furthermore define a group of 549 genes that are constitutively expressed (≤2.5% change in expression). Interestingly, hotspots, and in particular those occupied by all 15 factors, are highly enriched near induced genes compared with regions occupied by fewer factors (Figures 4E, S3D, and S3E). Although we cannot exclude the possibility that hotspots also play a role in transcriptional repression, these analyses strongly suggest that hotspots are key regulatory regions involved in activating the gene program associated with early adipocyte differentiation.

Hotspots Are Central Constituents in Super-Enhancers

Whyte et al. (2013) and Lovén et al. (2013) recently reported the existence of super-enhancers, which are large regulatory regions in the genome that have a high density of transcription factor binding sites and very high levels of MED1. To identify super-enhancers 4 hr after induction of differentiation of 3T3-L1 cells, we merged transcription factor binding sites in close proximity and defined super-enhancers and regular transcription factor binding regions based on the level of MED1 as shown in Figure 5A. Using this approach, we identified 340 super-enhancers that have ultrahigh levels of MED1 binding (Figure 5A) as well as p300 recruitment (Figure S4A). Because super-enhancers are composed of multiple constituent binding sites that were merged together in this analysis, they are much larger (median size of 33,740 bp) than regular transcription factor binding regions (Figure S4B). Importantly, we show that

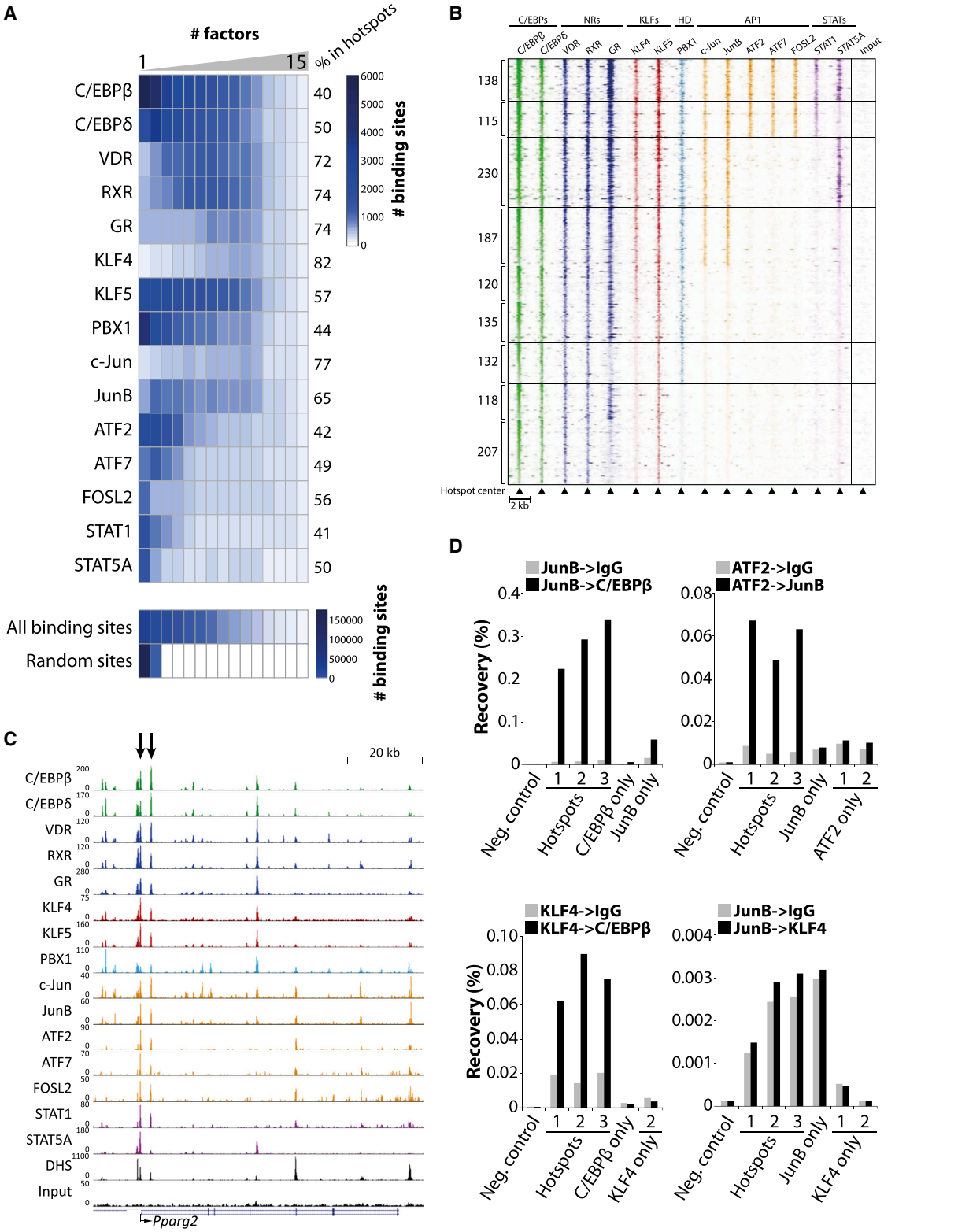
Figure 2. Combined Genomics and Proteomics Approach Reveals Candidate Regulators of Early Adipocyte Differentiation of 3T3-L1 Cells

(A) Main GO categories from the STRING database (Szklarczyk et al., 2011) associated with the C/EBPβ-interacting proteins identified by MS analysis of C/EBPβ immunoprecipitates on nuclear extract from 3T3-L1 cells induced to differentiate for 4 hr. All proteins were identified in two independent experiments, except for the transcription factors FOSL2, GR, STAT1, and PBX1, which were only identified in one of the replicates.

(B) All of the transcriptional regulators identified in (A) were subdivided into more specific GO categories, and the fold enrichment in the C/EBPβ IP relative to the control IP using a nonspecific IgG antibody is shown. Transcription factors that we subsequently subjected to genomics analyses using ChIP-seq are indicated by an asterisk.

(C) Summary of the results from motif analyses of DHS sites 4 hr after induction of differentiation (Siersbæk et al., 2011). Motifs were identified by MEME analysis (Bailey et al., 2009). The transcription factors identified in the C/EBPβ coIP that have been shown to bind to these motifs are indicated.

See also Figure S1 and Tables S1 and S2.



(legend on next page)

constituent binding regions within super-enhancers have higher levels of MED1 and transcription factor binding than regular transcription factor binding regions outside super-enhancers (Figures 5B and S4D). This suggests that constituents within super-enhancers cooperate to establish a large enhancer region comprised of multiple particularly strong enhancers. Taken together, these findings show that the extremely high levels of MED1 recruitment to super-enhancers reflect the facts that (1) super-enhancers are comprised of multiple individual binding regions, and (2) constituent binding regions in super-enhancers have on average much higher levels of MED1 recruitment compared with regular transcription factor binding regions, presumably at least in part as a consequence of high levels of transcription factor binding.

It is interesting to note that early hotspots both within and outside super-enhancers are enriched for PPARγ binding in mature adipocytes (Haakonsson et al., 2013; Figure S4C), indicating that a subset of early established hotspots may remain active enhancers also in mature adipocytes and be involved in regulating the mature adipocyte gene program. This is consistent with our previous finding that many chromatin regions that become accessible within the first 4 hr of differentiation remain open throughout the differentiation process (Siersbæk et al., 2011).

Interestingly, early super-enhancers are highly enriched near early-induced genes and depleted near early-repressed genes, whereas regular transcription factor binding regions (which also include hotspots) outside super-enhancers are not enriched near regulated genes (Figure 5C). Similarly, we identified super-enhancer-associated genes by assigning each region to the nearest gene. Consistent with the findings above, the super-enhancer-associated genes are induced during the first 4 hr of differentiation, whereas genes associated with regular transcription factor binding regions in general show no change in mRNA levels (Figure S4E). Super-enhancer-associated genes are enriched in GO terms linked to the early phase of the differentiation process, including extracellular matrix-receptor interactions, cell proliferation, and growth factor binding (Figure S4F). Thus, super-enhancers appear to be central drivers of the early transcriptional reprogramming that defines this phase of the differentiation process (examples of super-enhancer-associated genes are shown in Figure 5E). Intriguingly, transcription factor binding regions occupied by multiple transcription factors, in particular hotspots occupied by all 15 investigated factors, are highly enriched in super-enhancer regions compared with binding sites occupied by fewer factors (Figure 5D). In fact, practically all super-enhancers (99%) contain at least one hotspot. Taken together, these results suggest that hotspots are central constituents in super-enhancer regions that control the gene program that drives the early phase of the adipocyte differentiation process.

Effect of Transcription Factor Perturbation on Hotspot and Super-Enhancer Activity

To obtain functional insight into transcription factor cooperativity in the formation of hotspots and super-enhancers, we perturbed transcription factor activity using two different approaches and analyzed MED1 recruitment using ChIP-seq (Figure 6, left). First, we induced cells to differentiate for 4 hr using the normal adipogenic cocktail (i.e., fetal bovine serum, insulin, a cAMP-elevating agent, and dexamethasone) or the adipogenic cocktail without the strong GR agonist, dexamethasone. Second, we performed short hairpin RNA (shRNA)-mediated knockdown of C/EBPβ (Figure S4G) prior to induction of differentiation using the normal adipogenic cocktail.

Interestingly, omission of dexamethasone from the adipogenic cocktail had a significantly greater effect on MED1 recruitment to GR binding sites within super-enhancers compared with GR binding sites outside super-enhancers (Figure 6A, middle), suggesting that cooperation between constituents in super-enhancers is particularly sensitive to perturbation of GR. Consistent with the notion that constituents within super-enhancers cooperate to recruit coactivators, MED1 recruitment to constituent binding regions without GR in super-enhancers is also significantly more affected by omission of dexamethasone than is MED1 recruitment to binding sites outside super-enhancers. Remarkably, the effect of dexamethasone omission on MED1 binding is significantly less for hotspots, whether located in super-enhancers or not, compared with binding regions occupied by few factors (Figure 6A, right). Thus, hotspots are much less sensitive to GR perturbation than non-hotspots, which indicates that transcription factor cooperation at the level of hotspots can compensate for the loss of GR.

In contrast to these findings, knockdown of the general transcription factor C/EBPβ affects C/EBPβ binding sites within and outside super-enhancers to the same extent (Figure 6B, middle). Furthermore, the effect of C/EBPβ knockdown on

Figure 3. Transcription Factors in the Early Adipogenic Network Colocalize at Transcription Factor Hotspots

(A) For each factor, the number of binding sites that are occupied by one (only the factor itself) to 15 factors is shown along with the percentage of binding sites that are located in hotspot regions. The numbers of all transcription factor binding sites based on the ChIP-seq data sets and all binding sites redistributed randomly in the genome that are occupied by one to 15 factors are shown at the bottom.

(B) Heatmap of transcription factor binding in a 2 kb region around the center of the nine largest groups of hotspots. Input signal (Siersbæk et al., 2011) is shown as a control.

(C) Screen shot from the UCSC genome browser (http://genome.ucsc.edu; Kent et al., 2002) showing the binding profiles of 15 transcription factors as well as DHS-seq data (Siersbæk et al., 2014) and input control (Siersbæk et al., 2011) at the *Pparg2* locus 4 hr after induction of differentiation. The two arrows point to two hotspots occupied by all 15 investigated factors.

(D) Re-ChIP results for four different transcription factor pairs at three hotspots as well as control sites. Hotspots 1–3 refer to regions in the *Pparg2* promoter, downstream of *BC026439*, and in an *Xrcc4* intron, respectively (see Figure S2D). The negative control is not occupied by any of the investigated factors, whereas the control sites to the right in each subfigure are only occupied by one of the two factors investigated (screen shots for these regions are shown in Figure S2E). Results are representative of two independent experiments.

See also Figure S2.

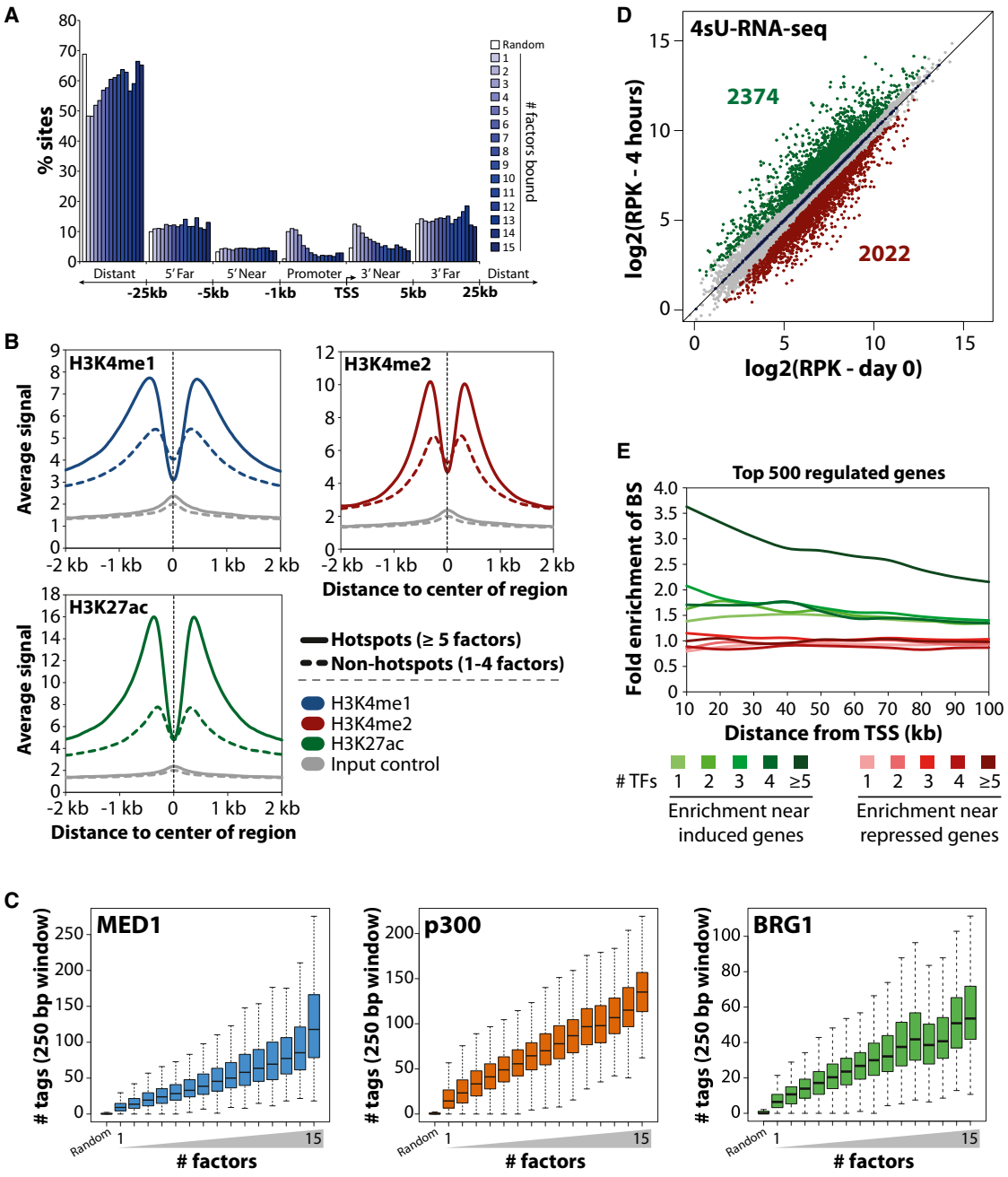


Figure 4. Extensive Transcription Factor Cooperativity at the Level of Hotspots

(A) Location of transcription factor binding sites occupied by one to 15 factors relative to the transcription start site (TSS) of RefSeq genes. The location of randomly placed binding sites of the same size is shown as a reference.

(B) The level of three histone marks characteristic of enhancers regions (i.e., H3K4me1, H3K4me2, and H3K27ac) in the vicinity of distal (>2 kb away from the TSS) non-hotspots (occupied by one to four factors) and hotspots (occupied by ≥5 factors). Input (Siersbæk et al., 2011) is shown as a control.

(C) Number of sequence tags at the regions defined in (A) for the Mediator subunit MED1, the histone acetyltransferase p300, and the SWI/SNF chromatin remodeling factor BRG1.

(D) Scatterplot showing the number of exon reads per kilobase (RPK) for all expressed genes (13,019). Significantly ($p \leq 0.01$) induced genes are green (2,374) and repressed genes are red (2,022). A group of constitutive genes (blue, 549 genes) was defined as those having ≤2.5% change in expression, and the rest of the nonregulated genes are colored gray.

(E) Enrichment of different types of binding sites (i.e., those occupied by one, two, three, or four factors, or at least five factors) near the top 500 most induced and repressed genes, respectively. Enrichment was determined as the number of binding sites per gene within different distances from the TSS (10–100 kb) of regulated genes relative to the number of binding sites per gene of constitutive genes as defined in (D).

See also Figure S3.

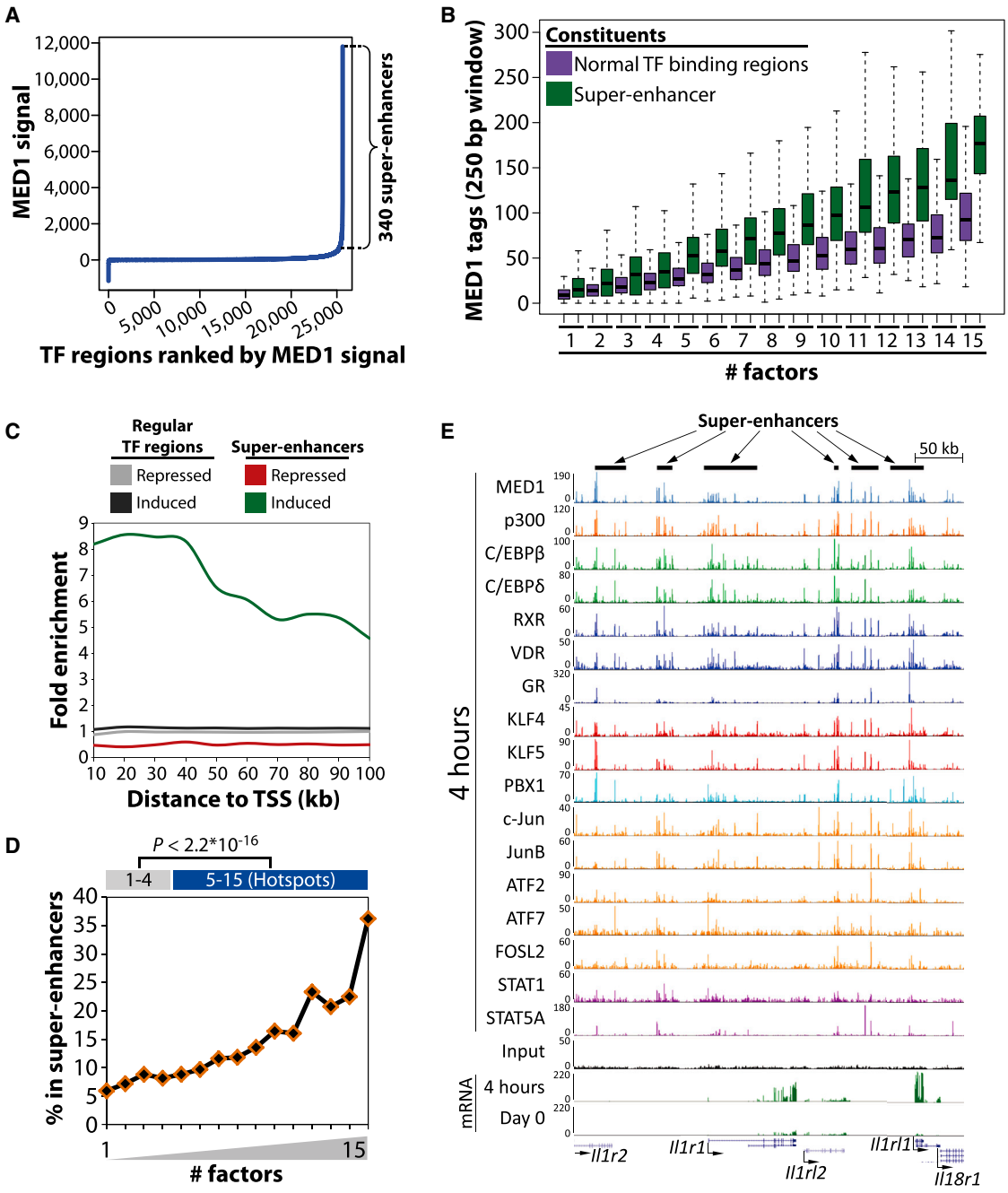


Figure 5. Hotspots Are Enriched in Early Adipogenic Super-Enhancers

(A) All the identified transcription factor binding sites (54,724) that were within 12.5 kb of each other were merged, resulting in 25,632 regions. These regions were ranked by their MED1 signal, where the input background (Siersbæk et al., 2011) had been subtracted. Regions with a MED1 signal (minus background) above 700 reads per 10 M total reads were defined as super-enhancers. All other regions were denoted as regular transcription factor binding regions.

(B) Number of MED1 sequence tags in constituents (250 bp window) in normal transcription factor binding regions and super-enhancers occupied by one to 15 factors.

(C) Enrichment of super-enhancers and regular transcription factor binding sites in the vicinity of the top 500 most regulated genes. Enrichment was determined as in Figure 4E.

(D) Fraction of transcription factor binding sites occupied by one to 15 factors that are found in super-enhancer regions. The significance of the higher occurrence of hotspots (i.e., binding sites occupied by at least five factors) relative to non-hotspots (i.e., binding sites occupied by one to four factors) within super-enhancer regions as determined by Fisher's exact test is shown at the top.

(E) Screen shot from the UCSC genome browser (http://genome.ucsc.edu; Kent et al., 2002) showing six super-enhancers in the vicinity of several genes, including *Il1r1* and *Il1rl1*, which are highly induced.

See also Figure S4.

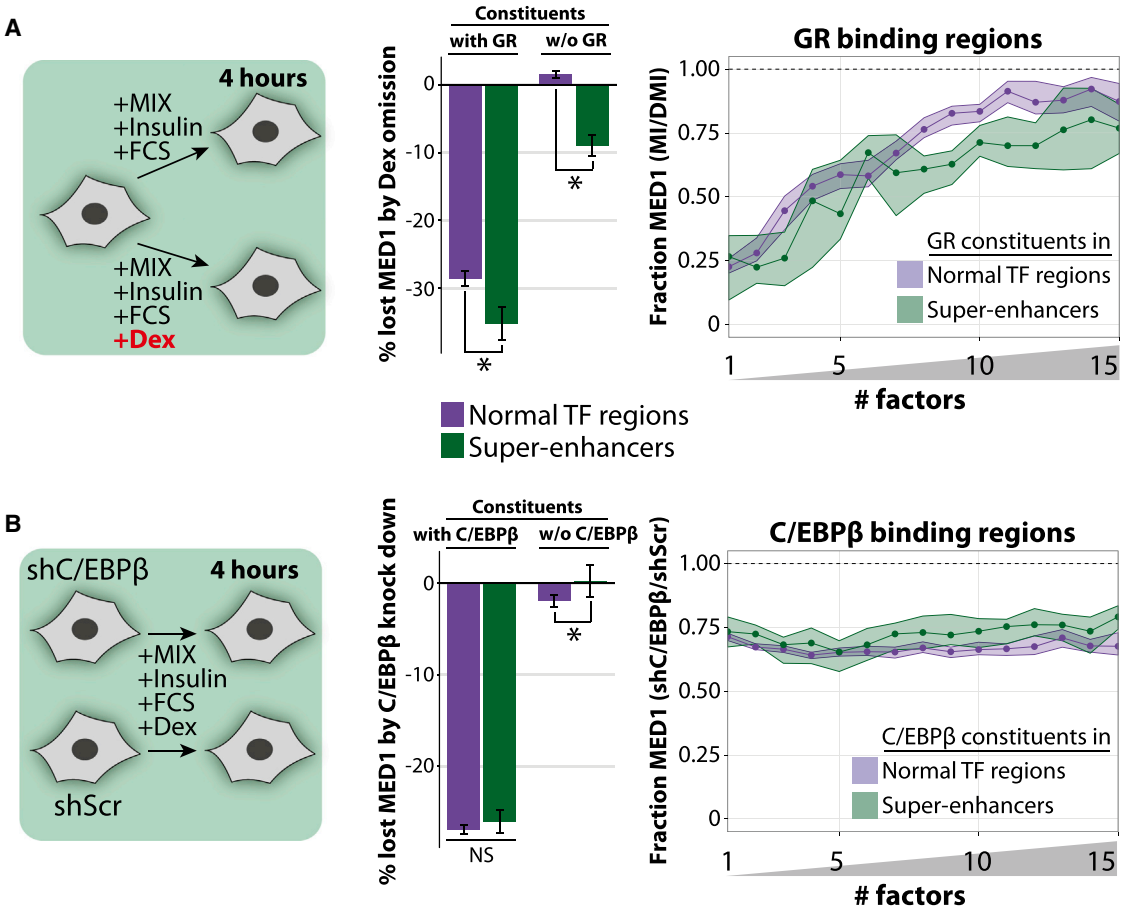


Figure 6. Transcription Factors Differ in their Relative Importance for Super- and Regular-Enhancer Activity

(A and B) Schematic illustration of the experimental setup (left). The mean loss of MED1 recruitment to constituents in regular transcription factor binding regions and super-enhancers upon omission of dexamethasone from the adipogenic cocktail or shRNA-mediated C/EBPβ knockdown is shown in the middle. Error bars illustrate the 95% confidence interval around the mean. To the right is shown the median fraction of MED1 recruitment retained upon omission of dexamethasone or knockdown of C/EBPβ, as described above at regions occupied by one to 15 factors within or outside super-enhancers. The transparent ribbon shows the 95% confidence interval around the median as determined by bootstrapping (Canty and Ripley, 2013; Davison and Hinkley, 1997). *$p < 0.05$, Student's t test. See also Figure S4.

MED1 recruitment is not influenced by the number of transcription factors recruited (Figure 6B, right). Taken together, these results indicate that although loss of GR can be compensated for by transcription factor cooperativity at the level of hotspots, the same is not the case for C/EBPβ, which is equally important for MED1 recruitment irrespectively of how many other factors are associated with a binding region. Furthermore, the results indicate that transcription factors differ in their relative importance for the activity of super-enhancers and regular transcription factor binding regions.

DISCUSSION

Here, we describe the relationship between super-enhancers and hotspots and take major steps toward understanding the complexity of both types of chromatin regions during genomic reprogramming associated with early adipogenesis. Motif analysis of DNA sequences at specific chromatin regions (e.g., DNase I hypersensitive sites or regions enriched for specific histone marks) is a commonly used strategy for identifying new candidate transcription factors involved in regulating a particular transcriptional response (Carroll et al., 2005; Heinz et al., 2010; Mikkelsen et al., 2010; Siersbæk et al., 2011; Steger et al., 2010). In general, however, motif searches have low specificity (Wasserman and Sandelin, 2004), and in addition, related transcription factors bind to very similar motifs (Sandelin and Wasserman, 2004), which makes it challenging to identify cognate transcription factors from a motif search alone. Here, we demonstrate that the combination of proteomics-based identification of proteins associated with a known key regulator of adipocyte differentiation (i.e., C/EBPβ) and motif analyses of accessible chromatin regions is a very powerful approach for identifying a large repertoire of factors in the complex transcription factor network that controls early adipocyte differentiation. This approach is likely to be widely applicable for investigating the transcription factor networks that control other biological processes.

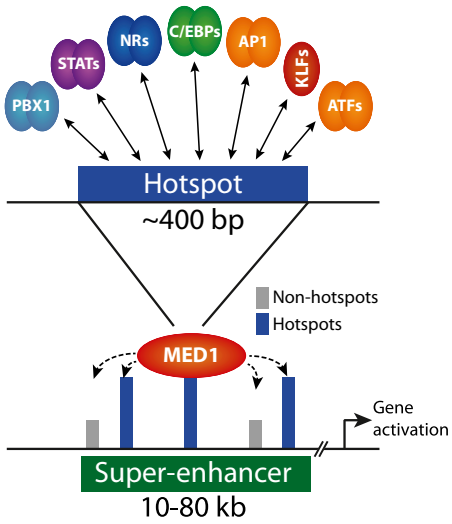


Figure 7. Model of Transcription Factor Cooperativity in Adipogenic Hotspots and Super-Enhancers

Multiple diverse transcription factors colocalize at small genomic regions termed transcription factor hotspots (~400 bp), which are central constituents in large super-enhancers (10–80 kb). Super-enhancers are characterized by very high levels of MED1 recruitment, and several lines of evidence suggest that constituents within super-enhancers cooperate to recruit MED1. Ultimately, establishment of super-enhancer regions results in activation of nearby genes characteristic of the early phase of adipogenesis.

One of the key findings from our study is that transcription factors, through the formation of hotspots, cooperate in recruitment of coactivators, chromatin remodeling, and establishment of an active epigenomic signature, as well as the activation of nearby genes during early adipogenesis. Thus, we suggest that crosstalk between transcription factors in hotspots is important for developmental reprogramming of the genome. Based on sequential ChIP experiments, we demonstrate that several transcription factor pairs associate simultaneously with hotspot regions, but, interestingly, we also find evidence of sequential and mutually exclusive binding of transcription factors to hotspots, indicating a dynamic exchange of factors at these sites. This is consistent with previous findings demonstrating transient interactions between transcription factors and chromatin (McNally et al., 2000; Métivier et al., 2003; Shang et al., 2000; Voss et al., 2011).

Importantly, our data indicate that cooperativity between transcription factors extends beyond hotspots to also include cooperativity between the constituent binding sites of super-enhancers (Figure 7), i.e., the large clusters of transcription factor binding regions that were recently reported to be central drivers of gene programs that define cell identity (Lovén et al., 2013; Whyte et al., 2013). We identify 340 super-enhancer regions that appear to be central drivers of the gene programs that are activated acutely (i.e., within 4 hr) by the adipogenic cocktail, and we show that hotspots are highly enriched in these super-enhancers (Figure 7). Our finding of higher levels of MED1 binding in super-enhancer constituents compared with regular transcription factor binding regions indicates that in addition to transcription factor cooperativity on a small genomic scale in hotspots, individual enhancers within super-enhancer regions also appear to cooperate to recruit coactivators, presumably through chromatin looping (Figure 7). Consistent with this, we find that perturbation of GR binding by omission of dexamethasone affects MED1 recruitment not only to super-enhancer constituents that bind GR but also to their neighboring constituents that do not bind GR.

Our finding that GR constituents in super-enhancers are more sensitive to perturbation of GR than GR-binding regions outside super-enhancers is consistent with the recent finding that super-enhancers are particularly sensitive to drug treatment in cancer cells (Lovén et al., 2013). This indicates that these adipogenic super-enhancers may present new drug targets for controlling adipocyte differentiation, which is of high clinical relevance. Interestingly and in contrast to what was found for GR, perturbation of C/EBPβ affects the activity of super-enhancers and regular binding regions to the same extent, demonstrating that transcription factors differ in their relative importance for the activity of super-enhancers and regular enhancers. This is consistent with a more general role of C/EBPβ in enhancer establishment.

In conclusion, we demonstrate extensive colocalization of transcription factors in hotspots that are important components of super-enhancers. Importantly, we show that transcription factor cooperativity plays a key role in defining enhancer activity at the level of hotspots as well as super-enhancers. These results indicate that hotspots and super-enhancers function as central hubs that serve to integrate external signals through transcription factor colocalization on chromatin.

EXPERIMENTAL PROCEDURES

Cell Culture

3T3-L1 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% calf serum. Cells were induced to differentiate in DMEM supplemented with 10% fetal bovine serum, 1 μM dexamethasone, 0.5 mM 3-isobutyl-1-methylxanthine, and 1 μg/ml insulin essentially as described previously (Helledie et al., 2002).

shRNA-Mediated Knockdown

Knockdown of C/EBPβ was performed essentially as described previously (Siersbæk et al., 2011). Briefly, 3T3-L1 cells were transduced with pSicoR PGK puro (12084; Addgene) lentivirus expressing shRNA against C/EBPβ or shRNA with a scrambled sequence at ~70% confluency in growth media supplemented with 8 μg/ml polybrene. Cells were then grown to confluence and induced to differentiate 2 days after reaching confluence as described above. MED1 ChIP-seq was performed for two independent biological replicates.

ChIP-Seq

ChIP was performed essentially as described previously (Siersbæk et al., 2012a). The following antibodies were used for immunoprecipitation: VDR (C-20, sc-1008; Santa Cruz), KLF4 (GKLF, H-180, sc-20691; Santa Cruz), c-Jun (H-79, sc-1694; Santa Cruz), PBX1 (Cat. No. 4342; Cell Signaling), KLF5 (a kind gift from Dr. Huck-Hui Ng), STAT1 (E-23, sc-346; Santa Cruz), JunB (210, sc-73; Santa Cruz), ATF2 (N-96, sc-6233; Santa Cruz), ATF7 (S-15, sc-19764; Santa Cruz), FOSL2 (L-15, sc-171; Santa Cruz), RXR (ΔN-197, sc-774; Santa Cruz), p300 (N-15, sc-584; Santa Cruz), MED1 (M-255, sc-8998; Santa Cruz), BRG1 (2822-1; Epitomics), H3K27ac (ab4729; Abcam), H3K4me1 (ab8895; Abcam), and H3K4me2 (9726; Cell Signaling). ChIP for all transcription factors and BRG1 was performed on formaldehyde crosslinked chromatin, and two

biological experiments were pooled. ChIP on H3K27ac, H3K4me1, and H3K4me2 was performed on two biological replicates of formaldehyde crosslinked chromatin and sequenced independently. ChIP for p300 and MED1 was performed once on chromatin that had been crosslinked in 2 mM disuccinimidyl glutarate (DSG) for 45 min and subsequently crosslinked by formaldehyde for 10 min. Chromatin-immunoprecipitated DNA was subjected to deep sequencing on the Illumina platform according to the instructions from the manufacturer (Nielsen and Mandrup, 2014). ChIP-seq data for C/EBPβ, C/EBPδ, GR, STAT5A, and input control were obtained from Siersbæk et al. (2011) and PPARγ ChIP-seq data were obtained from Haakonsson et al. (2013).

4sU-RNA-Seq

4sU-RNA-seq was performed essentially as described previously (Rabani et al., 2011). Briefly, cells were incubated with 400 μM 4sU for 30 min and then harvested in TRIzol (Invitrogen). RNA was then purified according to the manufacturer's instructions. Purified RNA was biotinylated and pulled down using streptavidin beads. Enriched RNA was then purified and subjected to standard mRNA sample preparation for sequencing, and sequenced on the Illumina platform according to the manufacturer's instructions.

Alignment, Peak Calling, and Gene Regulation Analyses

Sequence tags were aligned to the genome (mm9) using Bowtie (Langmead et al., 2009). ChIP-seq peaks were called using HOMER (Heinz et al., 2010), and regulated genes were identified using the DESeq package in R (Anders and Huber, 2010). Intersections between genomic position files were generated using BEDTools (Quinlan and Hall, 2010).

CoIP

Nuclei were isolated 4 hr after induction of 3T3-L1 differentiation and lysed by low-grade sonication. Nuclear extract was obtained by centrifugation (20,000 × *g*, 30 min, 4°C). Cleared nuclear extract was subjected to coIPs overnight at 4°C using C/EBPβ (sc-150 AC; Santa Cruz) or IgG control (sc-2345 AC; Santa Cruz) antibodies conjugated to agarose beads. After extensive washing, immunoprecipitates were eluted by boiling in SDS buffer and subsequently analyzed by MS/MS.

Nano-High-Performance Liquid Chromatography MS/MS Analysis of Immunoprecipitates

CoIP samples were subjected to in-gel digestion with trypsin and subsequently analyzed using an EasyLC nanoLC (Proxeon) coupled with an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific). MS/MS spectra were processed and analyzed by using Proteome Discoverer (v1.4.0.288; Thermo Fisher Scientific).

Additional information regarding the materials and methods used in this work is available in Supplemental Experimental Procedures.

ACCESSION NUMBERS

The GEO accession number for the sequencing data reported in this paper is GSE56872.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.04.042.

AUTHOR CONTRIBUTIONS

R.S., A.R., O.N.J., and S.M. conceived and designed the study. R.S., A.R., R.N., S.S., S.T., A.L., and L.L.C.P. performed the experiments. R.S. and A.R. analyzed the data. A.R.-W., O.N.J., and S.M. supervised the study. R.S. and S.M. wrote the manuscript with input from the other authors.

ACKNOWLEDGMENTS

The authors are grateful to members of the Mandrup and Jensen groups for valuable discussions. In particular, the authors thank Mads M. Aagaard, Anders K. Haakonsson, Søren F. Schmidt, and Jesper G. Madsen for assistance with bioinformatics analyses. We also thank Dr. Huck-Hui Ng for the generous gift of the KLF5 antibody. This work was in part carried out at the Villum Center for Bioanalytical Sciences, Department of Biochemistry and Molecular Biology, SDU, supported by the Villum Foundation. Work in the Mandrup laboratory was supported by grants from the Danish Independent Research Council | Natural Sciences and the Novo Nordisk Foundation, and work in the Jensen laboratory was financed by the Danish National Research Foundation (grant DNRF82 to the Center for Epigenetics).

Received: October 10, 2013
Revised: April 2, 2014
Accepted: April 18, 2014
Published: May 22, 2014

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. *37* (Web Server issue), W202–W208.

Biddie, S.C., John, S., Sabo, P.J., Thurman, R.E., Johnson, T.A., Schiltz, R.L., Miranda, T.B., Sung, M.-H., Trump, S., Lightman, S.L., et al. (2011). Transcription factor AP1 potentiates chromatin accessibility and glucocorticoid receptor binding. Mol. Cell *43*, 145–155.

Birsoy, K., Chen, Z., and Friedman, J. (2008). Transcriptional regulation of adipogenesis by KLF4. Cell Metab. *7*, 339–347.

Blumberg, J.M., Tzameli, I., Astapova, I., Lam, F.S., Flier, J.S., and Hollenberg, A.N. (2006). Complex role of the vitamin D receptor and its ligand in adipogenesis in 3T3-L1 cells. J. Biol. Chem. *281*, 11205–11213.

Boergesen, M., Pedersen, T.Å., Gross, B., van Heeringen, S.J., Hagenbeek, D., Bindesbøll, C., Caron, S., Lalloyer, F., Steffensen, K.R., Nebb, H.I., et al. (2012). Genome-wide profiling of liver X receptor, retinoid X receptor, and peroxisome proliferator-activated receptor α in mouse liver reveals extensive sharing of binding sites. Mol. Cell. Biol. *32*, 852–867.

Canty, A., and Ripley, B. (2013). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-9. http://cran.r-project.org/web/packages/boot/boot.pdf

Carroll, J.S., Liu, X.S., Brodsky, A.S., Li, W., Meyer, C.A., Szary, A.J., Eeckhoute, J., Shao, W., Hestermann, E.V., Geistlinger, T.R., et al. (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell *122*, 33–43.

Chang, C.J., Chen, Y.L., and Lee, S.C. (1998). Coactivator TIF1beta interacts with transcription factor C/EBPbeta and glucocorticoid receptor to induce alpha1-acid glycoprotein gene expression. Mol. Cell. Biol. *18*, 5880–5887.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Cianferotti, L., and Demay, M.B. (2007). VDR-mediated inhibition of DKK1 and SFRP2 suppresses adipogenic differentiation of murine bone marrow stromal cells. J. Cell. Biochem. *101*, 80–88.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.

Davison, A.C., and Hinkley, D.V. (1997). Bootstrap Methods and Their Applications (Cambridge, UK: Cambridge University Press).

Farmer, S.R. (2006). Transcriptional control of adipocyte formation. Cell Metab. *4*, 263–273.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

Grøntved, L., John, S., Baek, S., Liu, Y., Buckley, J.R., Vinson, C., Aguilera, G., and Hager, G.L. (2013). C/EBP maintains chromatin accessibility in liver and facilitates glucocorticoid receptor recruitment to steroid response elements. EMBO J. *32*, 1568–1583.

Haakonsson, A.K., Stahl Madsen, M., Nielsen, R., Sandelin, A., and Mandrup, S. (2013). Acute genome-wide effects of rosiglitazone on PPARγ transcriptional networks in adipocytes. Mol. Endocrinol. *27*, 1536–1549.

He, A., Kong, S.W., Ma, Q., and Pu, W.T. (2011). Co-occupancy by multiple cardiac transcription factors identifies transcriptional enhancers active in heart. Proc. Natl. Acad. Sci. USA *108*, 5632–5637.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Helledie, T., Grøntved, L., Jensen, S.S., Kiilerich, P., Rietveld, L., Albrektsen, T., Boysen, M.S., Nøhr, J., Larsen, L.K., Fleckner, J., et al. (2002). The gene encoding the Acyl-CoA-binding protein is activated by peroxisome proliferator-activated receptor gamma through an intronic response element functionally conserved between humans and rodents. J. Biol. Chem. *277*, 26821–26830.

Hurtado, A., Holmes, K.A., Ross-Innes, C.S., Schmidt, D., and Carroll, J.S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nat. Genet. *43*, 27–33.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lefterova, M.I., and Lazar, M.A. (2009). New developments in adipogenesis. Trends Endocrinol. Metab. *20*, 107–114.

Lefterova, M.I., Zhang, Y., Steger, D.J., Schupp, M., Schug, J., Cristancho, A., Feng, D., Zhuo, D., Stoeckert, C.J., Jr., Liu, X.S., and Lazar, M.A. (2008). PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. Genes Dev. *22*, 2941–2952.

Lovén, J., Hoke, H.A., Lin, C.Y., Lau, A., Orlando, D.A., Vakoc, C.R., Bradner, J.E., Lee, T.I., and Young, R.A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell *153*, 320–334.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958–970.

Mariani, O., Brennetot, C., Coindre, J.-M., Gruel, N., Ganem, C., Delattre, O., Stern, M.-H., and Aurias, A. (2007). JUN oncogene amplification and overexpression block adipocytic differentiation in highly aggressive sarcomas. Cancer Cell *11*, 361–374.

McNally, J.G., Müller, W.G., Walker, D., Wolford, R., and Hager, G.L. (2000). The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. Science *287*, 1262–1265.

Métivier, R., Penot, G., Hübner, M.R., Reid, G., Brand, H., Kos, M., and Gannon, F. (2003). Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter. Cell *115*, 751–763.

Mikkelsen, T.S., Xu, Z., Zhang, X., Wang, L., Gimble, J.M., Lander, E.S., and Rosen, E.D. (2010). Comparative epigenomic analysis of murine and human adipogenesis. Cell *143*, 156–169.

Mink, S., Haenig, B., and Klempnauer, K.H. (1997). Interaction and functional collaboration of p300 and C/EBPbeta. Mol. Cell. Biol. *17*, 6609–6617.

Monteiro, M.C., Sanyal, M., Cleary, M.L., Sengenès, C., Bouloumié, A., Dani, C., and Billon, N. (2011). PBX1: a novel stage-specific regulator of adipocyte development. Stem Cells *29*, 1837–1848.

Moorman, C., Sun, L.V., Wang, J., de Wit, E., Talhout, W., Ward, L.D., Greil, F., Lu, X.-J., White, K.P., Bussemaker, H.J., and van Steensel, B. (2006). Hotspots of transcription factor colocalization in the genome of Drosophila melanogaster. Proc. Natl. Acad. Sci. USA *103*, 12027–12032.

Nielsen, R., and Mandrup, S. (2014). Genome-wide profiling of transcription factor binding and epigenetic marks in adipocytes by ChIP-seq. Methods Enzymol. *537*, 261–279.

Nielsen, R., Pedersen, T.A., Hagenbeek, D., Moulos, P., Siersbaek, R., Megens, E., Denissov, S., Børgesen, M., Francoijs, K.-J., Mandrup, S., and Stunnenberg, H.G. (2008). Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. Genes Dev. *22*, 2953–2967.

Oishi, Y., Manabe, I., and Nagai, R. (2011). [Krüppel-like family of transcription factor 5 (KLF5). KLF5 is a key regulator of adipocyte differentiation]. Nippon Rinsho *69* (*Suppl 1*), 264–268.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., et al. (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. *29*, 436–442.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279–283.

Rosen, E.D., and MacDougald, O.A. (2006). Adipocyte differentiation from the inside out. Nat. Rev. Mol. Cell Biol. *7*, 885–896.

Sandelin, A., and Wasserman, W.W. (2004). Constrained binding site diversity within families of transcription factors enhances pattern discovery bioinformatics. J. Mol. Biol. *338*, 207–215.

Shang, Y., Hu, X., DiRenzo, J., Lazar, M.A., and Brown, M. (2000). Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. Cell *103*, 843–852.

Siersbæk, R., Nielsen, R., John, S., Sung, M.-H., Baek, S., Loft, A., Hager, G.L., and Mandrup, S. (2011). Extensive chromatin remodelling and establishment of transcription factor 'hotspots' during early adipogenesis. EMBO J. *30*, 1459–1472.

Siersbæk, M.S., Loft, A., Aagaard, M.M., Nielsen, R., Schmidt, S.F., Petrovic, N., Nedergaard, J., and Mandrup, S. (2012a). Genome-wide profiling of peroxisome proliferator-activated receptor γ in primary epididymal, inguinal, and brown adipocytes reveals depot-selective binding correlated with gene expression. Mol. Cell. Biol. *32*, 3452–3463.

Siersbæk, R., Nielsen, R., and Mandrup, S. (2012b). Transcriptional networks and chromatin remodeling controlling adipogenesis. Trends Endocrinol. Metab. *23*, 56–64.

Siersbæk, R., Baek, S., Rabiee, A., Nielsen, R., Traynor, S., Clark, N., Sandelin, A., Jensen, O.N., Sung, M.-H., Hager, G.L., et al. (2014). Molecular Architecture of Transcription Factor Hotspots in Early Adipogenesis. Cell Rep. *7*, this issue, 1434–1442.

Steger, D.J., Grant, G.R., Schupp, M., Tomaru, T., Lefterova, M.I., Schug, J., Manduchi, E., Stoeckert, C.J., Jr., and Lazar, M.A. (2010). Propagation of adipogenic signals through an epigenomic transition state. Genes Dev. *24*, 1035–1044.

Szklarczyk, D., Franceschini, A., Kuhn, M., Simonovic, M., Roth, A., Minguez, P., Doerks, T., Stark, M., Muller, J., Bork, P., et al. (2011). The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. *39* (Database issue), D561–D568.

Tanaka, T., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Defective adipocyte differentiation in mice lacking the C/EBPbeta and/or C/EBPdelta gene. EMBO J. *16*, 7432–7443.

Tang, Q.-Q., Otto, T.C., and Lane, M.D. (2003). CCAAT/enhancer-binding protein beta is required for mitotic clonal expansion during adipogenesis. Proc. Natl. Acad. Sci. USA *100*, 850–855.

Voss, T.C., Schiltz, R.L., Sung, M.-H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L. (2011). Dynamic exchange at regulatory elements during chromatin remodeling underlies assisted loading mechanism. Cell *146*, 544–554.

Wang, H., and Scott, R.E. (1994). Adipocyte differentiation selectively represses the serum inducibility of c-jun and junB by reversible transcription-dependent mechanisms. Proc. Natl. Acad. Sci. USA *91*, 4649–4653.

Wasserman, W.W., and Sandelin, A. (2004). Applied bioinformatics for the identification of regulatory elements. Nat. Rev. Genet. *5*, 276–287.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Zhang, J.-W., Tang, Q.-Q., Vinson, C., and Lane, M.D. (2004). Dominant-negative C/EBP disrupts mitotic clonal expansion and differentiation of 3T3-L1 preadipocytes. Proc. Natl. Acad. Sci. USA *101*, 43–47.

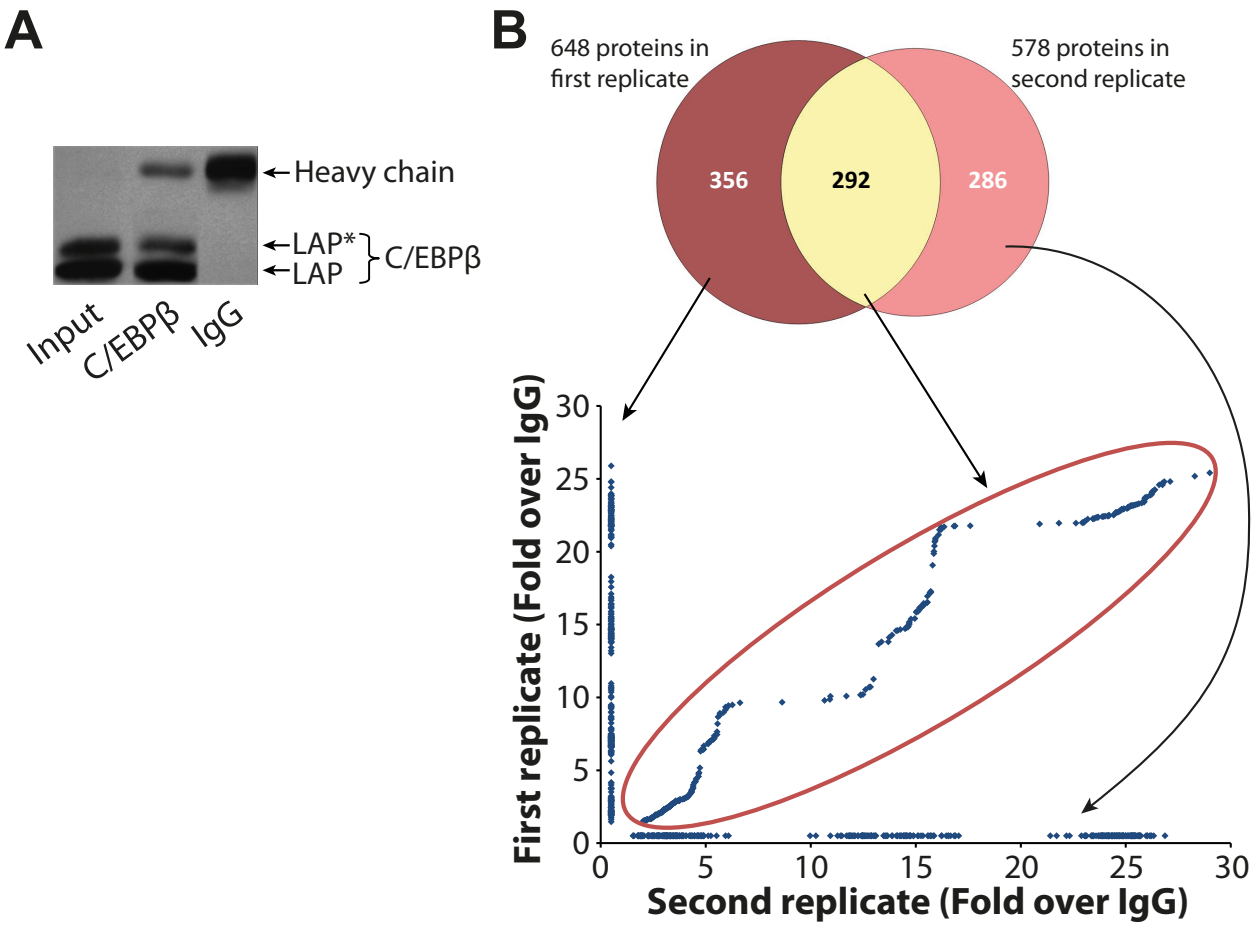
Cell Reports, Volume *7*

Supplemental Information

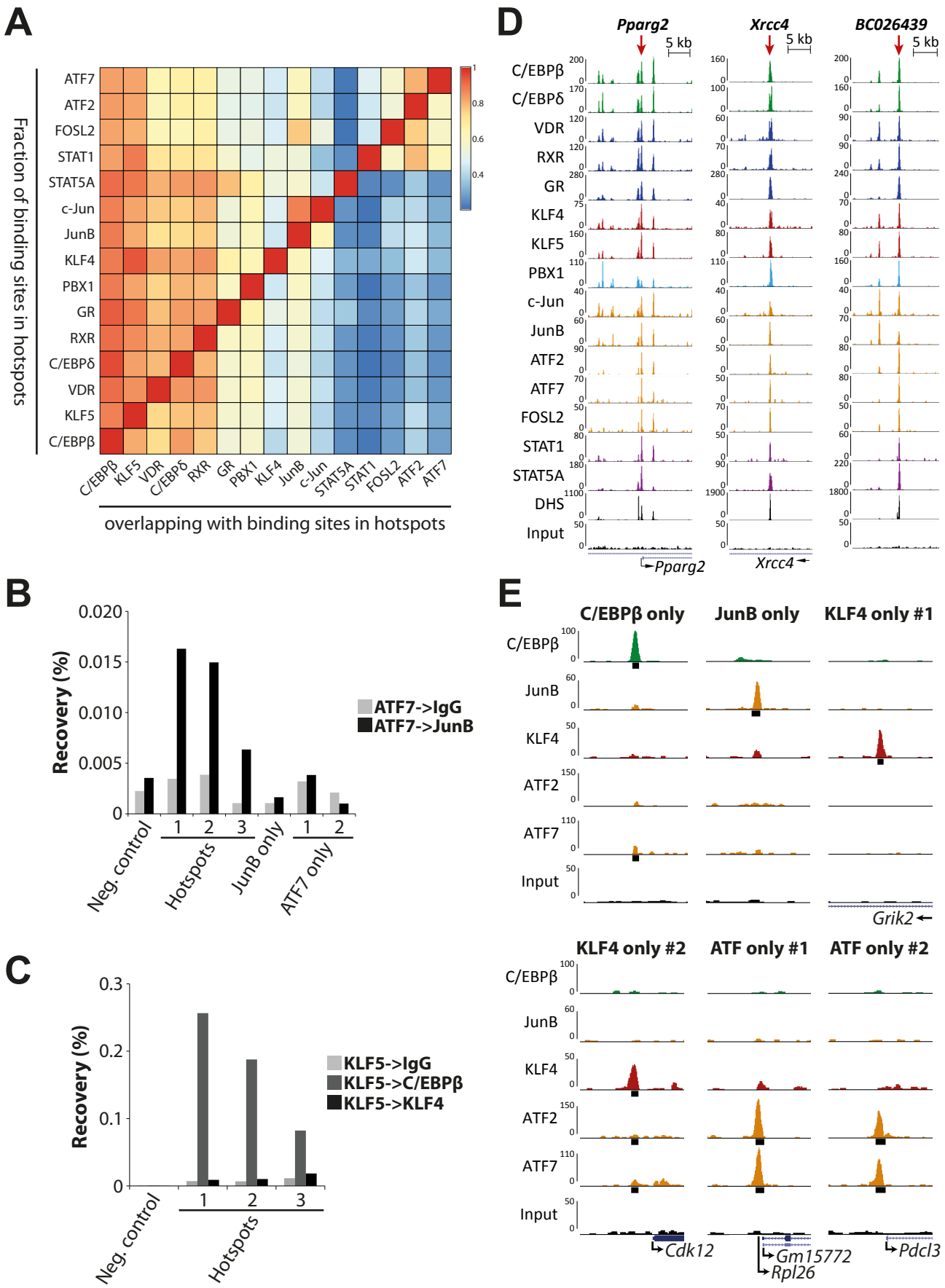
Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers

Rasmus Siersbæk, Atefeh Rabiee, Ronni Nielsen, Simone Sidoli, Sofie Traynor, Anne Loft, Lars La Cour Poulsen, Adelina Rogowska-Wrzesinska, Ole N. Jensen, and Susanne Mandrup

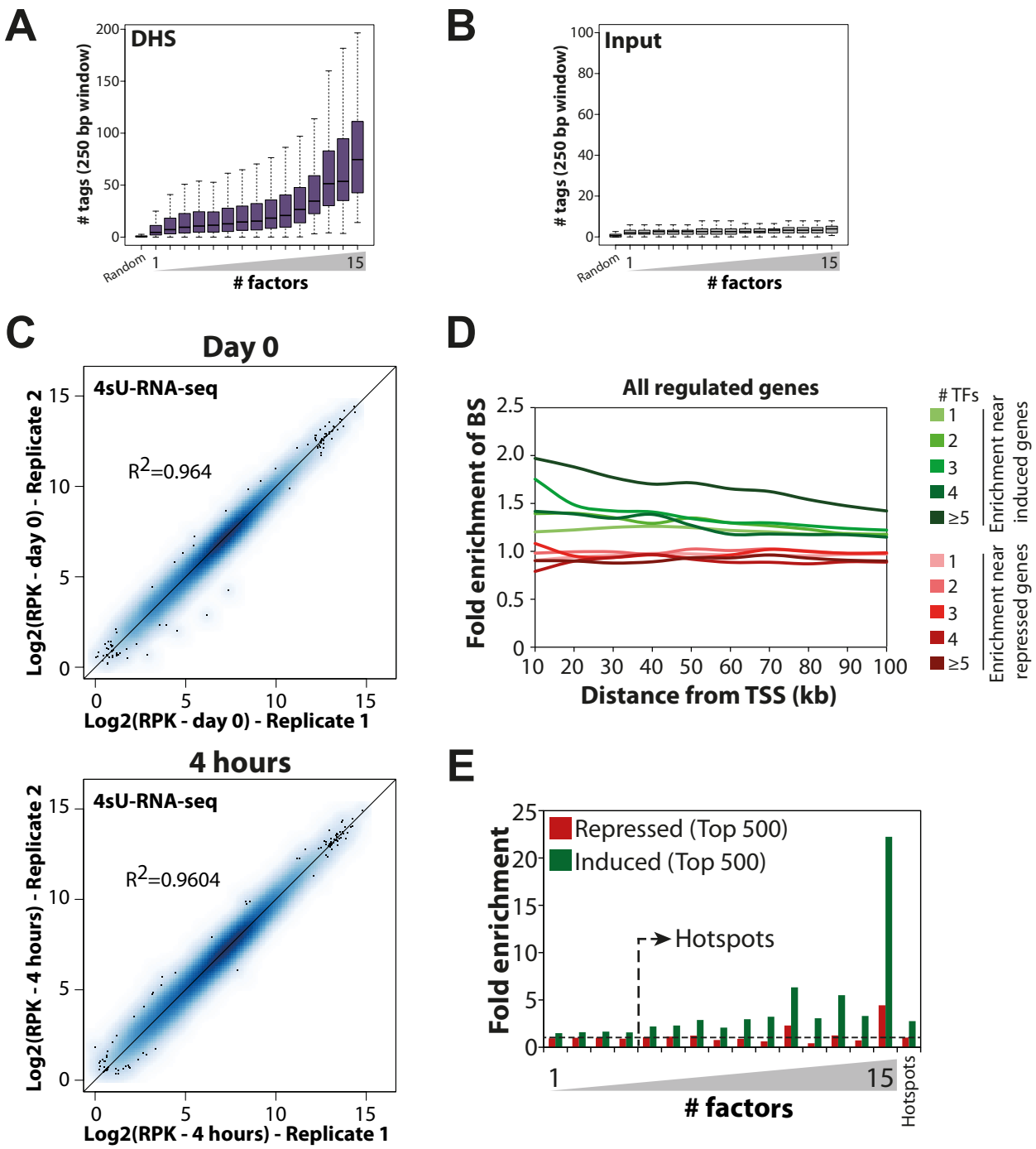
Figure_S1


A
Heavy chain
LAP*
LAP
C/EBPβ
Input
C/EBPβ
IgG
B
648 proteins in first replicate
578 proteins in second replicate
356
292
286
First replicate (Fold over IgG)
Second replicate (Fold over IgG)
0
5
10
15
20
25
30

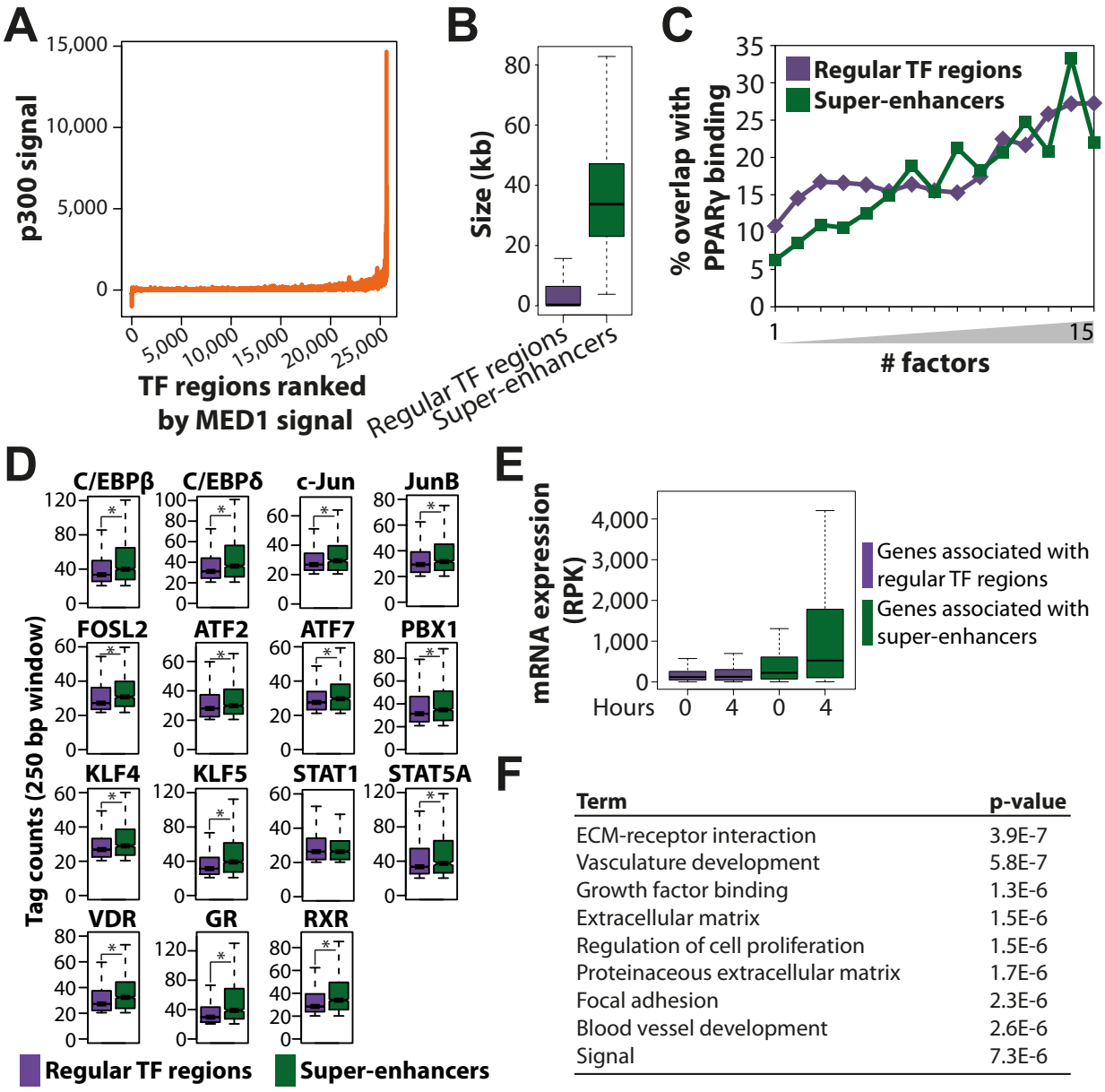
Figure_S2


A
Fraction of binding sites in hotspots
ATF7
ATF2
FOSL2
STAT1
STAT5A
c-Jun
JunB
KLF4
PBX1
GR
RXR
C/EBPδ
VDR
KLF5
C/EBPβ
overlapping with binding sites in hotspots
B
Recovery (%)
ATF7->IgG
ATF7->JunB
Neg. control
Hotspots
JunB only
ATF7 only
C
Recovery (%)
KLF5->IgG
KLF5->C/EBPβ
KLF5->KLF4
Neg. control
Hotspots
D
Pparg2
Xrcc4
BC026439
5 kb
C/EBPβ
C/EBPδ
VDR
RXR
GR
KLF4
KLF5
PBX1
c-Jun
JunB
ATF2
ATF7
FOSL2
STAT1
STAT5A
DHS
Input
Pparg2
Xrcc4
E
C/EBPβ only
JunB only
KLF4 only #1
KLF4 only #2
ATF only #1
ATF only #2
Grik2
Cdk12
Gm15772
Rpl26
Pdcl3

Figure_S3

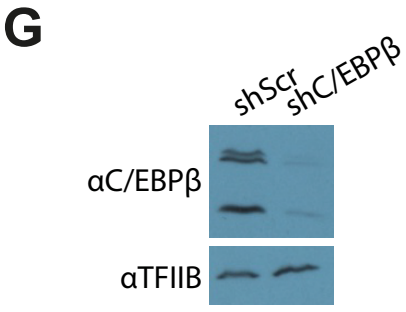

A
DHS
tags (250 bp window)
200
150
100
50
0
Random
1
15
factors
B
Input
tags (250 bp window)
100
80
60
40
20
0
Random
1
15
factors
C
Day 0
4sU-RNA-seq
R2=0.964
Log2(RPK - day 0) - Replicate 2
Log2(RPK - day 0) - Replicate 1
4 hours
4sU-RNA-seq
R2=0.9604
Log2(RPK - 4 hours) - Replicate 2
Log2(RPK - 4 hours) - Replicate 1
D
All regulated genes
Fold enrichment of BS
Distance from TSS (kb)
TFs
1
2
3
4
≥5
Enrichment near induced genes
Enrichment near repressed genes
E
Repressed (Top 500)
Induced (Top 500)
Hotspots
Fold enrichment
1
15
factors
Hotspots

Figure_S4



F

Term	p-value
ECM-receptor interaction	3.9E-7
Vasculature development	5.8E-7
Growth factor binding	1.3E-6
Extracellular matrix	1.5E-6
Regulation of cell proliferation	1.5E-6
Proteinaceous extracellular matrix	1.7E-6
Focal adhesion	2.3E-6
Blood vessel development	2.6E-6
Signal	7.3E-6



Supplemental Figure Legends

Figure S1. Identification of 292 C/EBPβ-associated proteins by mass-spectrometry based proteomics (Related to Figure 2). A) Western blot showing robust pull down of C/EBPβ specifically in the C/EBPβ immunoprecipitation and not in the control experiment using an IgG control antibody. B) Venn diagram showing overlap between the sets of proteins identified in two biological replicates of mass spectrometry analyses of C/EBPβ-associated proteins (top). Quantification results of the identified proteins are shown at the bottom.

Figure S2. Extensive co-localization of transcription factors at hotspots (Related to Figure 3). A) Fraction of binding sites for the factors on the y-axis overlapping with binding of the factors on the x-axis in hotspot regions. B,C) Re-ChIP results for ATF7-JunB, KLF5-KLF4, and KLF5-C/EBPβ as in Figure 3D. Results are representative of two independent experiments. Screen shots for the three hotspots and control regions are shown in panels D and E. D) Screen shots from the UCSC genome browser (http://genome.ucsc.edu) (Kent et al., 2002) showing DHS-seq (Siersbæk et al., 2014, this issue of Cell Reports) and ChIP-seq data at hotspots used for the re-ChIP experiments shown in Figure 3D and in panels B and C of this figure. E) Screen shots from the UCSC genome browser showing ChIP-seq data for the binding regions used as control regions in Figure 3D and panels B and C of this figure.

Figure S3. Hotspos are key regulators of gene activation (Related to Figure 4). A,B) Number of DHS-seq (Siersbæk et al. 2014, this issue of Cell Reports) (A) and input control (Siersbæk et al., 2011) (B) sequence tags at the regions defined in Figure 4C. C) Scatterplot showing the number of exon reads per kilobase (RPK) for the two 4sU-RNA-seq replicates generated in preadipocytes before induction of differentiation (day 0) and four hours after induction of differentiation. The R^2 values are shown in each plot. D) Enrichment of different types of transcription factor binding sites in the vicinity of all regulated genes as defined in Figure 4D. Enrichment was determined as in Figure 4E. E) Enrichment of binding regions

occupied by 1 through 15 factors within 50 kb of the top 500 regulated genes. Enrichment was determined as in Figure 4E.

Figure S4. Super-enhancers are large genomic regions associated with gene activation (Related to Figure 5 and 6). A) p300 signal at the merged transcription factor binding sites ranked by MED1 as in Figure 5A. B) Size of super-enhancers and regular transcription factor binding regions outside super-enhancers. C) The fraction of constituents in normal transcription factor binding regions and super-enhancers occupied by 1 through 15 factors that become occupied by PPARγ in mature adipocytes. PPARγ ChIP-seq data was obtained from Haakonsson et al., 2013. D) Boxplots show the binding intensity (i.e. number of ChIP-seq reads) at constituents in normal transcription factor binding regions and super-enhancers that are occupied by each factor. E) mRNA expression during the first four hours of differentiation shown as exon reads per kilobase (RPK) for genes associated with super-enhancers and for those associated with regular transcription factor binding regions outside super-enhancers. Transcription factor binding regions were assigned to the nearest gene in these analyses. F) Significantly enriched biological terms found to be associated with super-enhancers. The super-enhancer associated genes were analyzed using DAVID (Huang et al., 2009a, b) to identify significantly enriched terms. G) Western blot showing C/EBPβ protein levels in 3T3-L1 cells transduced with lentivirus expressing scrambled shRNA (shScr) or shRNA targeting C/EBPβ (shC/EBPβ) after four hours of differentiation. TFIIB is shown as a loading control.

Tables S1 and S2

Data is supplied as separate excel-files. Both tables relate to Figure 2.

Table S1. List of C/EBPβ-associated proteins identified by mass spectrometry-based analyses (first replicate). Co-immunoprecipitation (co-IP) was performed on nuclear extract from 3T3-L1 cells induced to differentiate for four hours using antibodies against C/EBPβ (sample) and IgG (control). Proteins specifically pulled down in the C/EBPβ IP were identified by mass spectrometry. The table shows the number of identified peptides and the ratio between the signal in the C/EBPβ IP and the signal in the IgG control IP.

Table S2. List of C/EBPβ-associated proteins identified by mass spectrometry-based analyses (second replicate). Data were generated as in Table S1 on a second independent biological replicate.

Supplemental Experimental Procedures

Alignment and analyses of ChIP-seq data

Sequence tags for ChIP-seq experiments were aligned to the mouse genome (mm9) using Bowtie (Langmead et al., 2009) with the following options: --best --strata -m 3. Reads that align to the exact same location in the genome were discarded. Aligned reads were analyzed by HOMER (Heinz et al., 2010) to identify peaks of tag enrichment. Peaks were called for each transcription factor if they were four times enriched relative to the local background (20 kb around the peak), four times enriched relative to a 3T3-L1 input control sample, and significant at a false discovery rate of 0.01%. A master set of peaks containing all the identified peaks for all factors (if the center of two or more peaks were within 200 bp, the peaks were merged) was defined. All peaks from this master set containing at least 20 tags per 10M tags in a 250 bp window around the center of each peak for a given transcription factor were called as high-confidence transcription factor binding sites. HOMER was also used for counting tags at the identified binding sites, generating matrices for heat maps, and generating average tag counts. BEDTools (Quinlan and Hall, 2010) was used to analyze position information for ChIP-seq peaks (e.g. intersection of peak files).

Alignment and analyses of 4sU-RNA-seq data

Reads from 4sU-RNA were aligned to the genome using an approach similar to that described previously (Habegger et al., 2011). Briefly, tags were aligned to the mm9 genome and to a pseudo genome containing all possible exon-exon junctions for all genes using Bowtie (Langmead et al., 2009) with the following options: --best --strata –m 1. Mapped tags were subsequently combined. Tags within exons were determined using HOMER (Heinz et al., 2010), and the tag counts were then analyzed using the DESeq package in R (Anders and Huber, 2010). Significantly regulated genes were called using an adjusted *P*-value of 0.01.

Nuclear extract

3T3-L1 fibroblasts were differentiated for four hours by stimulation with 3-isobutyl-1-methylxanthine, dexamethasone, and insulin as described previously (Helledie et al., 2002). Approximately 10^8 cells were used for nuclei isolation. After rinsing twice with cold PBS, cells were scrapped off the dish in cold PBS containing protease inhibitors (Complete, EDTA-free Protease Inhibitor cocktail tablets, Roche Applied Science (#05056489001)) and phosphatase inhibitors (1 mM sodium ortho-vanadate (NaOVan), 10 mM NaF, 10 mM β-glycerol-phosphate) on ice. After centrifugation at 1,200xg for 5 min at 4 ^{o}C, the pellet was washed rapidly using 5X column volumes (CV) cold buffer A containing protease and phosphatase inhibitors (PIs) (20 mM Tris-HCl pH=7.5, 20 mM NaCl, 0.2 mM EDTA pH 8.0, PIs) and the supernatant was removed. The cell pellet was then resuspended in 5XCV of cold buffer A+PIs and kept on ice for 10 min in order to swell. After 10 min, 1 volume of cold buffer A+NP-40+PIs (20 mM Tris-HCl pH=7.5, 20 mM NaCl, 0.2 mM EDTA pH 8.0, 0.04% NP40, PIs) was added to the cell suspension, mixed gently and kept on ice for 10 min. Nuclei were spun down at 1,200xg for 7 min at 4 ^{o}C and the supernatant (cytosolic and cell membrane proteins and particles) was removed. The nuclear pellet was washed twice with 5XCV of cold buffer A+PIs, and the supernatant was removed each time after 5 min centrifugation at 4 ^{o}C. The nuclear pellet was then resuspended in 3XCV of cold buffer B+NP-40+PIs (20 mM Tris-HCl pH=7.5, 150 mM NaCl, 0.2 mM EDTA pH 8.0, 1% NP40, PIs) and kept on ice for 10 min. The nuclei were lysed by 6 cycles of sonication at low intensity using a Bioruptor (Diagenode). Cleared nuclear extract was obtained by 30 min centrifugation at 20,000xg, 4 ^{o}C. The cleared nuclear extract was divided into two equal amounts for immunoprecipitation (before division a part of nuclei extract was taken out and saved as input for further investigation by western blotting).

Co-Immunoprecipitation

Co-immunoprecipitation (Co-IP) was performed at 4 ^{o}C unless otherwise indicated, using C/EBPβ antibody conjugated to protein A/G agarose beads (sc-150 AC, Santa Cruz) in parallel with IgG antibody conjugated to protein A/G agarose beads (sc-2345 AC, Santa Cruz). The antibody-conjugated beads were incubated

with nuclear extract overnight at 4 ^{o}C on a rotator. Beads were washed three times in cold buffer B+1% NP-40+PIs and subsequently three times with cold buffer B+PIs without NP-40 (20 mM Tris-HCl pH=7.5, 150 mM NaCl, 0.2 mM EDTA pH 8.0). The immunoprecipitates were eluted with 150 ul of elution buffer (50 mM Tris-HCl pH=6.8, 2% SDS (BDH grade), 8% Glycerol, 0.008% bromophenol blue) at 100 ^{o}C for 5 min. Eluates were concentrated to 60 µl using vivaspin centrifugal concentrators 500 (VS0191) (10 ul of the original samples were saved for the further investigation by western blotting).

SDS-PAGE and in-gel digestion

Samples were run on a 4-12% Bis-Tris gel (NuPage) according to the manufacturer's instructions. The proteins were visualized with coomassie staining and the gel background was destained with water. The entire gel was diced into small pieces (1–2 mm), the gel pieces were then washed with water, and shrunk in acetonitrile. The gel pieces were subjected to in-gel proteins reduction and alkylation with 10 mM DTT and 20 mM iodoacetamide, respectively. After sequential washing and dehydration with water and 50% ACN, gel pieces were dried and rehydrated with 12.5 ng/µl trypsin solution in 50 mM ammonium bicarbonate on ice for 45 min. The digestion was continued at 37 ^{o}C overnight. The tryptic peptides were extracted with 5% formic acid/50% acetonitrile and concentrated with vacuum centrifugation (Anal. Chem. 1996, 68, 850-858). Samples were resuspended in 0.1% formic acid (solvent A).

Nano-HPLC MS/MS Analysis

Nano-LC-ESI-MS/MS was performed on an LTQ-Orbitrap XL mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with an EasyLC nanoLC (Proxeon, Odense, Denmark). The mobile phases consisted of 0.1% formic acid (solvent A) and 0.1% formic acid in 95% ACN (solvent B). Peptides were loaded onto a 20x0.1 mm trap column and separated by a 17 cm 75 µm ID picofrit column, both in-house packed with C18 resin (Reprosil-Pur C18-AQ 3 µm; Dr. Maisch GmbH, Germany). Flowrate of the analysis was 300 nL/min; with a linear gradient of solvent B from 5 to 35% over 45 min or 60 min. Spray voltage was 2.2 kV in positive-ion mode. Full scan spectra from m/z 300 to 1700 at resolution of 60,000 were acquired

in the orbitrap. The top seven most intense ions were selected each cycle for MS/MS fragmentation in the LTQ-XL ion trap using CID with normalized collision energy of 35. Dynamic exclusion was used with the following parameters: exclusion time 60 s, repeat count 1, repeat duration 1 s, exclusion mass width 10 ppm, and exclusion size 500. Singly charged species were excluded from MS/MS selection.

The acquired MS/MS spectra were analyzed with Proteome Discoverer (v1.4.0.288, Thermo Fisher Scientific). Mascot (v2.3.2, Matrix Science, London, UK) was chosen as search engine, and data were searched against the mouse Uniprot database (updated January 2013). Trypsin was selected as enzyme with two missed cleavage allowed. Precursor mass tolerance was 8 ppm, product mass tolerance was 0.6 Da. Fixed modification was carbamidomethyl (C). Dynamic modifications were: oxidation (M), deamidation (NQ) and acetylation (protein N-terminal). Results were filtered at a 5% false discovery rate (FDR), Mascot score 20, and peptide rank 1. Label-free quantification was included in the workflow by adding the node Precursor Area Detector in Proteome Discoverer, where chromatographic peak areas were integrated with a tolerance of 2 ppm. Statistics was performed for both biological experiments to obtain the relative enrichment values of quantified proteins in the C/EBPβ pull down sample compared to the control IgG pull down.

Western Blotting

Nuclear extract input and the C/EBPβ and IgG co-IP samples or whole cell extracts from 3T3-L1 cells transduced with lentivirus expressing scrambled shRNA or shRNA against C/EBPβ were submitted to SDS-PAGE and separated proteins were subsequently blotted onto a PVDF membrane. After blocking in milk, the membrane was then incubated with a primary antibody against C/EBPβ (sc-150, Santa Cruz) or TFIIB (sc-225, Santa Cruz) followed by an anti-rabbit peroxidase-conjugated secondary antibody (Dako) and developed by enhanced chemiluminescence (ECL).

Sequential ChIP (Re-ChIP)

Cross-linked chromatin from 3T3-L1 cells differentiated for four hours was immunoprecipitated with antibody against the first factor of interest as described in the main paper except that chromatin was eluted in a solution of 10 mM DTT and 1% SDS gently rotating for 30 minutes at room temperature. Eluted chromatin was diluted 20-fold, divided into two equal amounts, and subjected to the second immunoprecipitation with antibody against the second factor of interest as well as an IgG control. The second elution was done using a standard elution buffer (1% SDS, 0.1 M $NaHCO_3$ gently rotating for 30 minutes at room temperature). The isolated chromatin was decross-linked, purified by phenol-chloroform extraction, and analyzed using qPCR. Primers are available upon request.

Supplemental References

Habegger, L., Sboner, A., Gianoulis, T., Rozowsky, J., Agarwal, A., Snyder, M., and Gerstein, M. (2011). RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. Bioinformatics (Oxford, England) *27*, 281-283.

Huang, D.W., Sherman, B., and Lempicki, R. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

Huang, D.W., Sherman, B., and Lempicki, R. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.